MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 3, 2018/2019

DPS5018 – INTRODUCTION TO PROBABILITY & STATICTICS

(For DIT students only)

31 MAY 2019
09.00 am – 11.00 am
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **6** pages (**3** pages with **4** questions and **3** pages of Appendix). Key formulae are given in the Appendix.

2. Answer **ALL** questions.

3. Write your answers in the answer booklet provided.

4. All necessary working steps must be shown.

QUESTION 1

a) The following probability distribution for the random variable X is given in the table below.

x	1	2	3	4	5	6	7
$P(X=x)$	0.3	0.06	0.05	0.10	0.08	0.14	0.27

i) Calculate the expected value. (2 marks)

ii) Compute the probability of $P(2 < X \leq 5)$. (1.5 marks)

b) From a sample of 13 diploma students, 84.7% are planning to continue their degree in MMU.

i) Calculate the probability that all diploma students continue their degree in MMU. (1.5 marks)

ii) How many of the students will be continuing their degree in MMU? (1 mark)

c) Dr. Cinta consults an average of two patients a week.

i) What is the probability that he has no patient in a week? (1.5 marks)

ii) How many patients consulted him in a month? (0.5 mark)

iii) Find the probability that he has five patients and more in a particular month. (1.5 marks)

d) The price of holiday ground trip package to Krabi, Thailand follows a normal distribution with a mean of RM850 and a standard deviation of RM75. Find the probability that a randomly selected holiday ground trip package to Krabi, Thailand is less than RM910. (2.5 marks)

[TOTAL 12 MARKS]

QUESTION 2

a) In a random survey, 27 out of 126 students agree on embedding an online learning medium into statistics subject.

i) Find the point estimate for proportion of students agree on embedding an online learning medium into statistics subject. (1 mark)

ii) Construct a 95% confidence interval for the population proportion of those who agree on embedding an online learning medium into statistics subject. (4 marks)

Continued...

b) Suppose that we want to estimate the true proportion of malfunction in a very large shipment of smartphones, and we want to be at least 95% confident that the error is at most 0.03. How large a sample will we need if

i) the true proportion is not known? (2.5 marks)

ii) the true proportion does not exceed 0.13? (2.5 marks)

[TOTAL 10 MARKS]

QUESTION 3

a) In order to investigate the relationship between mean job tenure in years among workers who have a diploma certification or higher and those who do not, random samples of each type of worker were taken, with the following results. A researcher claims that the mean job tenure among those with diploma certification or higher education is greater than among those without diploma certification.

	n	$\bar{x}$	s
Diploma or higher	155	5.2	1.3
Without diploma	210	5.0	1.5

i) State the null and alternative hypotheses. (1 mark)

ii) Calculate the test statistic and p-value by using the hypothesis testing for two population means. (2.5 marks)

iii) At the 1% significance level, do the data provide sufficient evidence to conclude that the researcher's claim is true? (2.5 marks)

b) Five years ago, 3.9% of children in a certain region in Malaysia lived with someone other than a parent. Sociologist wishes to test whether the current proportion is different. Perform the relevant test at the 5% level of significance using a random sample of 2,759 children, found that by 139 children do not live with their own parent. (5 marks)

Continued...

c) The price of a popular tennis racket at a national chain store is RM179. Mary bought five of the same racket at different online stores for the following prices:

155	179	175	161	175

Determine whether there is sufficient evidence in the sample, at the 5% level of significance, to conclude that the average price of the racket is less than RM179 if purchased at an online store. Given that the sample mean and sample variance are 169 and 108 respectively. (4 marks)

[TOTAL 15 MARKS]

QUESTION 4

Michael wants to get a clear view regarding the apartment that he wants to buy. After collecting the information from n house developers, he managed to get the below results where x is the apartment size (in square feet) and y is the price of the apartment (in RM '000).

$SS_{xy} = 369600$	$\sum x = 17400$
$SS_{xx} = 1893470$	$\sum y = 1698$

a) Given $\sum xy = 4062750$, find the value of n. (2.5 marks)

b) By using the above information, determine the least square regression line, $\hat{y} = b_0 + b_1 x$. (2.5 marks)

c) Use the result in part 4(b) to estimate the price of the apartment if the size of the apartment is 1941 square feet. (1.5 marks)

d) Given $SS_{yy} = 91080$, calculate the correlation coefficient, r and interpret the answer. (1.5 marks)

e) At 5% significance level, do the data provide sufficient evidence to conclude that the apartment size and the price of the apartment are positively linearly correlated? (5 marks)

[TOTAL 13 MARKS]

Continued...

APPENDIX – KEY FORMULA

DISCRETE RANDOM VARIABLE, BINOMIAL DISTRIBUTION, AND POISSON DISTRIBUTION

- Mean of a discrete random variable x : $\mu = \sum xP(X = x)$
- Standard deviation of a discrete random variable x :

$$\sigma = \sqrt{\sum x^2 P(X = x) - \mu^2}$$

- Binomial probability formula : $P(X = x) = \binom{n}{x} p^x q^{n-x}$
- Poisson probability formula : $P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$
- Key formula for Binomial Probability Distribution:

If $p \le 0.5$	If $p > 0.5$
i) $P(X = r) = B(r) - B(r-1)$	i) $P(X = r) = B(n-r) - B(n-r-1)$
ii) $P(X \ge r) = 1 - B(r-1)$	ii) $P(X \ge r) = B(n-r)$
iii) $P(X \le r) = B(r)$	iii) $P(X \le r) = 1 - B(n-r-1)$
iv) $P(X > r) = 1 - B(r)$	iv) $P(a \le X \le b) = B(n-a) - B(n-b-1)$
v) $P(X < r) = B(r-1)$	
vi) $P(a \le X \le b) = B(b) - B(a-1)$	

- Key formula for Poisson Probability Distribution:
  - If $\mu = \lambda$

  i) $P(X = r) = Poi(r) - Poi(r-1)$
  ii) $P(X \ge r) = 1 - Poi(r-1)$
  iii) $P(X \le r) = Poi(r)$
  iv) $P(X > r) = 1 - Poi(r)$
  v) $P(X < r) = Poi(r-1)$
  vi) $P(a \le X \le b) = Poi(b) - Poi(a-1)$

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \dfrac{x - \mu}{\sigma}$

Continued...

ESTIMATION

- The $(1-\alpha)100\%$ confidence interval for μ is

$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$	where $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$
$\bar{x} \pm t_{\alpha/2}s_{\bar{x}}$	where $s_{\bar{x}} = \dfrac{s}{\sqrt{n}}$

- The $(1-\alpha)100\%$ confidence interval for p is

$\hat{p} \pm z_{\alpha/2}s_{\hat{p}}$	where $s_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

- **Sample size**

$n = \left[\dfrac{\left(Z_{\alpha/2}\right)\sigma}{E}\right]^2$	$n = \dfrac{\left(Z_{\alpha/2}\right)^2 pq}{E^2}$

HYPOTHESIS TESTING

Population	Mean	Proportion
ONE	$Z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ **or** $t = \dfrac{\bar{x}-\mu}{s/\sqrt{n}}$	$Z = \dfrac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}}$
TWO	$Z = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$Z = \dfrac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ where $\hat{p}_p = \dfrac{x_1+x_2}{n_1+n_2}$

Continued...

LINEAR REGRESSION ANALYSIS

- **Least square regression equation,** $\hat{y} = \beta_0 + \beta_1 x$ where $\beta_1 = \dfrac{SS_{xy}}{SS_{xx}}$ and

  $\beta_0 = \dfrac{1}{n}\left[\sum y - \beta_1 \sum x\right]$

- **Correlation coefficient,** $r = \dfrac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$ where

$SS_{xx} = \sum x^2 - \dfrac{\left(\sum x\right)^2}{n}$	$SS_{yy} = \sum y^2 - \dfrac{\left(\sum y\right)^2}{n}$	$SS_{xy} = \sum xy - \dfrac{\left(\sum x\right)\left(\sum y\right)}{n}$

- **Regression *t*-Test**

  $t = \dfrac{b_1 - \beta_1}{s_b}$ where $s_b = \dfrac{s_e}{\sqrt{SS_{xx}}}$ and $s_e = \sqrt{\dfrac{SS_{yy} - b_1 SS_{xy}}{n-2}}$

- **Correlation *t*-Test**

  $t = \dfrac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}$

End of Page.